OSM/高德路网匹配融合技术在道路空间化中的应用 *

王　康，朱欣焰，呙　维，盛光晓

(武汉大学 测绘遥感信息工程国家重点实验室，武汉 430079)

摘　要：“六合一”道路编码是交管业务中用来定位事故和违法的基础文本数据，缺乏空间位置信息，而已有的常用路网数据如高德路网，都是基于多车道路段表达的路网且现势性相对于 OSM 路网较低，难以满足交管业务的需求。针对上述问题，以高德路网作为基础、高现势性的 OSM（OpenStreetMap）路网做补充，将轨迹聚类分析中的 LCSS(longest common subsequence)算法应用在路网匹配过程中，并对匹配后的路网使用 Stroke 方法进行路网融合。实验结果表明，使用 LCSS 算法可以达到良好的路网匹配效果。最后，基于此开发了一套路网匹配融合程序，并在武汉市交通管理局投入使用。

关键词：路网匹配；LCSS 算法；路网融合

中图分类号：TP391　　doi: 10.3969/j.issn.1001-3695.2018.05.0392

Application of technology of OSM/Gaode road network's matching and fusion in road spatialization

Wang Kang, Zhu Xinyan[1], Guo Wei, Sheng Guangxiao

(*State Key Laboratory for Information Engineering in Surveying, Mapping & Remote Sensing, Wuhan University, Wuhan* 430079, *China*)

Abstract: The "Six-in-one" road code is the basic text data used to locate accidents and violations in the traffic management business and it lacks spatial information while existing road network data in common use, such as Gaode road network, is a kind of road network where one complete road is expressed by multiple road sections. Gaode road network'currency is relatively low compared to the OSM road network and it is hard to meet the demand for traffic management business. In order to solve the problems as above, this paper with Gaode Road Network as a basis and a high-currency OSM (OpenStreetMap) road network as a complement applied LCSS (Longest Common Subsequence) algorithm commonly used in trajectory clustering to road matching, then applied the stroke method in the road network fusion after the road network matching. The experiment results show that LCSS algorithm has a good effect in road network's matching. Finally, this paper developed a road network's matching and fusion program based on this, and it has been applied in the Wuhan Traffic Management Bureau.

Key words: road network matching; LCSS algorithm; road network fusion

0　引言

目前武汉市交通管理局在事故和违法的定位主要依靠的是“六合一”道路编码，该编码属于文本数据，缺乏空间信息，难以满足交管业务信息化、精细化的要求。而武汉市交通管理局现有的路网数据主要来源于高德多车道路网数据，具有数据更新不及时的问题。而且对于交管业务中的事故、违法定位，高德路网这种过于精细化的多车道路网，在实际应用中并不能满足要求，因此可以结合高现势性的 OSM 路网数据进行路网匹配，对“六合一”道路编码进行空间化。

空间目标匹配分为点要素匹配、线要素匹配、面要素匹配，而路网匹配属于线要素匹配。对于路网匹配方面的研究和应用，一直都是 GIS 领域最活跃的研究方向之一。目前线要素匹配的方法主要有以下五种：a）几何匹配，通过计算两个空间实体的几何相似度来进行匹配；b）拓扑匹配，通过计算同名实体的拓扑关系作为匹配依据；c）语义匹配，通过计算候选实体的语义名的相似度进行匹配；d）基于概率的匹配算法，通过计算实体匹配概率大小来确定匹配实体；e）智能算法，如蚁群算法、隐

收稿日期：2018-05-16；**修回日期：**2018-07-09　**基金项目：**国家重点研发计划课题（2018YFB0505503）；测绘遥感信息工程国家重点实验室重点开放基金资助项目；测绘遥感信息工程国家重点实验室专项科研经费资助项目

作者简介：王康（1993-），男，硕士研究生，主要研究方向为时空大数据分析、数据挖掘（1134314305@qq.com）；朱欣焰（1963-），男，教授，博士，主要研究方向为空间信息服务、空间数据库等；呙维（1981-），男，副教授，主要研究方向为三维建模；盛光晓（1993-），男，主要研究方向为交通 GIS、交通路网简化等.

马尔可夫模型，以此寻求全局最优的路网匹配方案。另外不少学者为了改进算法匹配准确度，会考虑加入约束条件或者结合上面多种算法来进行路网匹配。国内学者刘一宁[1]以上海市某一区域不同时相的道路作为同源数据，利用缓冲区增长算法进行路网匹配。郭庆胜等人[2]分别选取了比例尺变化较小的南昌市路网数据和比例尺变化较大的合肥市路网数据，提出了一种顾及尺度变化和数据更新的道路网匹配算法并进行路网匹配实验。张云菲等人[3]利用概率松弛法分别在武汉市和瑞士苏黎世两个地区进行多源路网匹配。巩现勇等人[4]实验证明利用蚁群算法寻求全局最优的道路网同名实体匹配方案是有效可行的。但是道路匹配研究大多都是同源路网匹配或者多尺度[5,6]的不同比例尺下的路网匹配。

本文基于武汉交管业务的需求，以高德路网数据为基础，利用现势性较好的 OSM 路网数据对武汉市路网数据（高德路网） 做补充，填补高德路网的部分数据的缺失，实现了异源路网的匹配。其中采用轨迹聚类算法中常用的 LCSS (longest common subsequence) [7] 算法将高德路网和 OSM（OpenStreetMap）路网做路网匹配。最后，在此基础上开发了一套路网匹配融合程序。

1　路网数据分析

本文采用的空间数据主要是高德路网数据和 OSM 路网数据。

1.1　高德路网数据

高德路网依赖高德地图先进的数据采集手段、自主研发的电子地图制作工艺和地理数据编辑、检核系统，形成了广泛、准确的地理信息。其数据质量较高，且绝大部分为自采数据。

高德路网数据较为完整，制作流程更为完善，几何精度更高，有着其他路网数据源无可比拟的优点。其坐标系采用的是一套经过加密的坐标系，称为火星坐标系，与 WGS84 坐标有着非线性的偏差。本文的高德路网数据囊括了武汉市 16 万余条的道路路段数据。

主要包括了表 1 所示的重要字段信息。

表 1　高德路网数据结构表

字段名	字段含义
MESH	道路所在图幅号
ROAD _ID	道路编号，道路在相应图幅内的编号
FNODE _ID	道路起始点在相应图幅内的编号
TNODE _ID	道路结束点在相应图幅号内的编号
DIRECTION	1 代表双向通行；2 代表正向通行；3 代表逆向通行；4 代表双向禁行
NAME _CHN	中文名称
ROAD_CLASS	道路等级，分为高速公路、国道、城市快速路、城市主干路等
LINK_TYPE	默认值为 0；1 代表轮渡航线；2 代表隧道；3 代表桥；4 代表地下通道
STATUS	默认是 0（正常）、1 代表建设中；2 代表禁行
OWNER_SHIP	默认为 0（公共道路）；1 代表内部道路；2 代表私有道路

1.2　OSM 路网数据

近十年来，基于用户创建地理数据的平台逐渐地兴起，各种类型(地图、卫星影像、路网数据、GPS 终端设备)的地理数据被用户收集并共享到 OpenStreetMap(OSM) 等应用平台中[8,9]。

OSM 自 2004 年上线以来，目前已拥有 200 万用户。因为操作简单，无须复杂的 GIS 技能，全球的地理信息爱好者就可以及时上传更新自己熟悉的道路信息。OSM 数据具有更新速度快、现势性高和获取成本低等优点。OSM 的这些优点已经引起了国内外 GIS 学者在 VGI 数据质量及可靠性、VGI 的应用等方面展开广泛的研究。比如 Zielstra 等人[10]对比了 OSM 和 TeleAtlas 在德国地区的数据，结果表明尽管 OSM 在郊区的数据覆盖率很小，但是在稍大点的城市（如 Berlin、Frankfurt、Munich），OSM 数据很丰富，可以作为许多项目的备用选择；罗路长等人[8]的实验结果表明，国内 OSM 数据的道路长度和道路名称属性完整性与 Baidu 地图、Google 地图相当；在韩国首尔，OSM 数据的道路长度和道路名称属性完整性与 Baidu 地图高，但比 Google 地图略低。Wang 等人[11]通过数据完整性、属性准确性、位置准确性这三个方面研究了武汉地区的 OSM 路网数据的质量，结果显示，OSM 数据可以作为一种新的城市高等级基础交通网络和郊区路网的数据来源。

武汉市的 OSM 路网有 1.9 万条数据，该路网数据主要包括表 2 所示的几个字段信息。

表 2　OSM 路网数据结构表

字段名	字段含义
Type	道路类型
Name	道路名
Tunnel	是否是隧道，0 代表不是，1 代表是
Bridge	是否是桥梁，0 代表不是，1 代表是
Oneway	是否是单向，0 代表不是，1 代表是

高德路网数据在与 OSM 路网数据进行匹配前，由于高德路网的精细化表达，主要存在以下四个难点：

a）高德路网一条完整的路被分割为多条路段，因此在匹配时需要根据空间上的衔接，需要判断是否为同一条道路。

b）高德路网的数据包含多车道信息，但是在交管事故、违法定位业务中，并没有区分道路的双向多车道问题，因此本文在处理的过程中需要对道路网进行路网简化，合并多车道，以及对相应的复杂线型的道路进行简化。在道路简化的过程中还会存在很多难点，下文会一一阐述。

c）路网信息缺失。高德路网虽然完整，但是仍然有接近一半的路网没有名称。

d)有 1 万多条路段存在别名，这部分路段都有自己的名字，这些路段的别名可能也是某些更宽泛概念上的某一段路，如 XX 国/省/县/乡道、二/三环线、绕城高速。

2　路网数据预处理

本文以高德路网数据为基础，在进行路网匹配前需要对路网数据预处理。高德路网作为多车道路网数据，可以确定的是高德路网线型复杂性是接下来进行路网数据处理的难点问题，而 OSM 路网为单车道路网，不需要进行预处理。在分析了高德路网数据的空间特点之后，主要考虑从五个方面对高德路网进行简化，即消除多车道间的小碎段、消除高速路上的匝道、多车道合并、消除主干道附近的碎段、复杂线型道路的简化。

2.1　消除多车道间的小碎段

在高德路网数据中，如图 1 所示的小碎段，在交管业务中，如事故定位等方面并不会有很大作用，只会增加道路网结构的复杂性，大大增加数据预处理的难度，在多车道合并和匝道消除的过程中产生较大干扰。

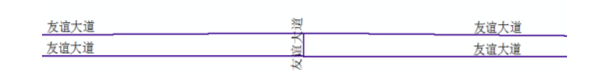


图 1　多车道间小碎段

因此，本文需要针对小碎段的特点作相应的简化处理。本文通过分析已有的武汉市高德路网的数据，得出小碎段的主要特点包括长度较短、与小碎段相邻或相交的同名路段、与小碎段构成的夹角较大。

针对上述特点，本文设定一个长度阈值 l 和一个角度阈值 θ，然后根据下面的处理流程进行处理：

a）筛选长度小于 l 的路段（选取合适的 $\boldsymbol{l}$ 值）

b）对 a）中符合条件的路段，逐条进行相交空间查询，找出每条路段有多少条路段与其相交。如果存在相交关系的同名路段数大于等于 2，并且都与该路段的夹角大于角度阈值 θ，就可以判定该路段为碎段，并予以剔除。

需要注意的是，本文提到的参数，如果没有特别说明，指的都是经验值。根据数据的不同，参数值视具体情况而定。

2.2　消除高速路的匝道

如图 2 所示，高速路匝道的存在使得路网的空间关系变得极为复杂。匝道作为高速公路的一种附属，而且部分匝道形状复杂，对数据预处理而言及其不友好。考虑到本文所做的研究工作主要在于交通领域的违法、事故定位等方面，因此道路层次关系和精细化的道路表达并不适用，所以对高速路进行了简化。

本文总结了匝道具有以下三个特点：

a）匝道长度较短；

b）匝道弯曲程度较大，呈弧线状；

c）一般与两条道路相交。

设定一个长度阈值、一个线段平均转角阈值和一个匝道子线段阈值。其中长度阈值指的是匝道的长度上限，线段平均转角阈值指的就是匝道本身转过的角度与匝道包含的子线段数（节点数减 1）的比值。选取长度小于长度阈值的路段。对于候选路段，对其进行空间分析，选取与之相接的路段。如果这些路段的不重名路段数大于等于 2，平均转角大于平均转角阈值且匝道子线段数大于匝道子线段阈值数，那么就可以断定该小碎段为高速路的匝道。数据处理的效果取决于长度阈值和角度阈值的选取，对于未能消除的匝道，适当进行手动处理。

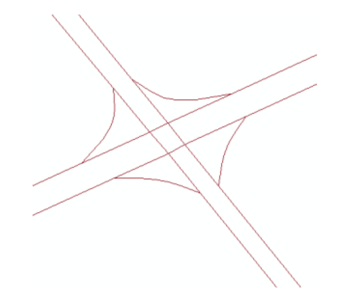


图 2　匝道分布

2.3　多车道合并

这一过程是数据预处理中涉及数据最多的一步。本文首先计算路段的斜率 k；然后在该路段的四分之一处、中心点、四分之三处，分别作斜率为-1/k 长度为 l 的线段；然后进行空间求交查询，若三者均存在与该线段相交且道路名称和选定路段相同的路段，且方向与选定路段较为接近，则可判定为是多余的车道，再根据多条车道求取中心线即为最终的简化后的道路。

经过多车道合并之后，道路网的结构得到了一定程度的简化和优化，但还是有一些情况无法得到处理，比如：

a）在路段长度较短且多车道之间的对应不够整齐的情况下，能被监测到可能性显著降低。

b）存在一定的误判，比如当两条路之间的间隔很小，但并不是一条多车道的路，这种情况一般存在于居民区的小巷。

为了改进数据处理的效果，结合一些先验知识来优化。针对情况 a)，通过相邻的同名道路方向转折角不大这一特点，将其作为一个标识器，判定路段的性质。针对情况 b)，通过该路段周围的路网特点（如果是小巷，就会在空间上具有网格状的特征模式），判定是同一条路还是两条相距较小的路。

2.4　消除主干道附近的碎段

路网中存在许多与主干道路相连或主干道的支路延伸出去的碎段，这部分碎段会影响路网的复杂性，删除这些不重要的碎段，对路网的完整性和结构影响不大。

2.5　复杂线型道路简化

部分道路线型复杂，通常是由上面几种形式组合而成，比如道路既是多车道，又是带有匝道的高速路，还包含部分碎段。对于这种情况就需要结合上述几种情形综合处理。

3　道路匹配与融合

3.1　路网匹配

高德路网中有近一半的道路道路名称是缺失的，如果不加以利用，会造成路网的空间数据一定程度的浪费,缺失的这部分道路在别的路网下可能存在相关的补充信息。而 OSM 作为一个开源的地图社区，聚集全世界的地图制作爱好者，数据积累

时间长，具有高效的地图更新频率，同时覆盖面广，可以作为一种可靠的地图数据源。本文使用 OSM 路网数据作为高德路网数据的补充，填补了部分高德路网数据路名信息的空白。因为匹配的数据源两方分别是高德地图数据和 OSM 路网数据，前者道路名缺失，所以只能通过空间数据的一些几何特性进行匹配。线性要素的几何特性主要包括位置、方向、长度、转动的角度（该线状要素总共转动了多少角度）。还有拓扑特性，即该要素与其他相邻要素之间的位置关系。本文采用的匹配方法主要利用位置、长度和方向信息。对于利用拓扑关系进行匹配这一点，考虑到两方数据源条目数多，且需要自建拓扑关系，因此在利用拓扑关系作为匹配条件，计算量太大，代价太高，不适合实际生产。

由于 OSM 路网数据和高德道路路网数据分别采用两种不同的坐标系，所以需要将 WGS84 坐标首先转换为高德火星坐标系。

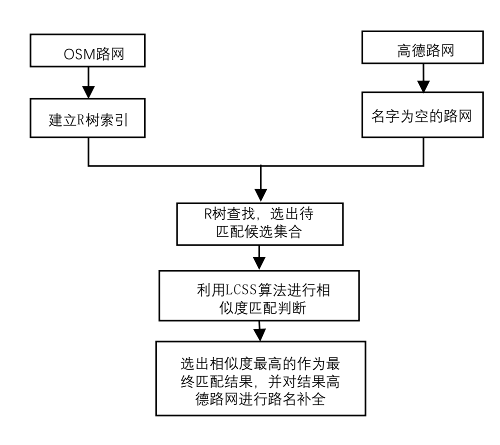


图 3　路网匹配流程

另外，两者数据条目大，其中 OSM 道路数据有一个特点：道路完整性更好，道路碎段较少。高德路网有 16W+数据量，OSM 路网有 1.9W 数据量。因此，为了加快匹配的处理速度，本文利用 R 树对待匹配路网 OSM 路网建立索引，路网匹配流程如图 3 所示。

在道路匹配的过程中，本文采用了 LCSS 算法进行道路的匹配。该部分内容在 3.1.2 中详述。

3.1.1 路网匹配方法概述

路网匹配是多源地图空间数据匹配融合中非常重要的一部分。从目前国内外的现有研究现状来看，基于几何特性的道路路网数据的匹配方法大体上分为以下两类：

a）缓冲区增长法[12,13]。该算法首先利用待匹配路段的缓冲区确定候选弧段集合，然后再直接对候选弧段进行匹配判断。如图 4 所示，合适的 Buffer-v 和 Buffer-p 数值对匹配结果更有利，一般会根据匹配情况自适应调整 Buffer-v 和 Buffer-p 的数值。初始 Buffer-v 和 Buffer-p 的数值是由经验决定的。

b）迭代最近点法（iterative closest point，ICP）[14]。具体思路是首先对道路网中弧段相交的节点进行初始匹配；根据节点匹配结果，继续对匹配节点相关联的弧段进行匹配判断。按照上面的步骤不断迭代循环，直到所有道路弧段的匹配全部结束。

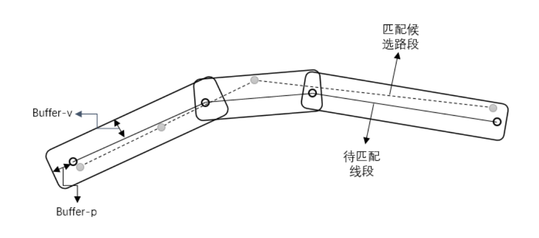


图 4　缓冲区增长法示意图

上面的两种方法都需要考虑路网中的几何特征。常见的几何特征包括距离、形状、角度等。道路弧段之间的距离计算不同于两点之间欧氏距离的计算，其计算过程较为复杂。常用的线要素距离测度分为两类：

a)基于弯曲的距离(Warping-based distance)。

欧氏距离、曼哈顿距离或者其他的 L-p 范数距离是最常用的距离。已有几个基于弯曲的距离方法分别是动态时间规整（dynamic time warping，DTW）[15]和之后的最长公共子序列（longest common subsequence，LCSS）[7]、编辑距离（edit distance on real sequence，EDR）。这些距离计算的方法都是同种定义，只是采用了不同的代价函数。

b)基于形状的距离(shape-based distance)。

主要有 Hausdorff 距离、Fréchet 距离。传统的 Hausdorff 距离在计算上容易实现，但是抗噪性较差。 为了解决原始 Hausdorff 距离在受到噪声污染或存在遮挡等情况，Huttenlocher（1993）提出了部分 Hausdorff 距离的概念[16]。另外一种度量距离是 Fréchet 距离，是法国数学家 Maurice René Fréchet 在 1906 年提出的一种路径空间相似性描述，具体原理是利用折线的顶点信息来近似计算曲线的 Fréchet 距离，其计算过程比较复杂。

3.1.2 LCSS 算法

鉴于 LCSS 算法常用于轨迹聚类分析中，较少应用在路网匹配中，本文尝试将其应用在路网匹配过程中，采用 LCSS 算法时，考虑到该算法有三个优势：

a)部分元素可以不必匹配到，在欧氏距离和DWT方法中，必须所有元素都进行匹配，即使是离群值。而本文在进行 OSM 路网和高德路网匹配时，两者的采样率不一样（OSM 路网由互联网用户标注，采样率低；高德路网由专业地图厂商完成数据采集，采样率更高），而且 OSM 路网基本保持了道路的完整性，高德路网对路网进行了分段处理，一条路可能被分割成多条路段。DWT 路网在处理长度不相同的数据时需要考虑截断，否则容易导致两者匹配程度下降；而 LCSS 算法不需要截断操作。因此，LCSS 算法正好满足了这种场景。

b) 根据已有的研究结果，相对于欧氏距离和 DTW，LCSS 模型计算更高效[7]。

c)欧氏距离和 DTW 这两种度量相似度的方法，抗噪性较差。但是由于 LCSS 算法不需要考虑到所有的数据，所以抗噪性更好。

LCSS（最长公共子序列）[7,17,18]是指两个或者多个序列中存在的最长的共同子序列。它的主要思想就是利用动态规划的思想计算出两个序列的最长公共子序列。但是在路网中，因为多源数据的差异性，两条道路不可能完全重合，所以判断是否拥有公共点的条件就是判断两个点的横纵坐标差值是否在阈值范围内。如果在阈值范围内，就可以看做是同一个点，则 LCSS 长度加 1。

首先，本文给出几个定义。假设所有的点都是二维平面上的点。A 和 B 是两条道路，点数分别为 n 和 m。$A_m=((a_{x,1},a_{y,1}),\ldots,(a_{x,m},a_{y,m}))$并且 $B_n=((b_{x,1},b_{y,1}),\ldots,(b_{x,n},b_{y,n}))$。

定义 1　LCSS 的具体定义如下：

下面的 LCSS (i,j) 表示 A 中的前 i 个点组成的子序列坐标串和 B 中的前 j 个点组成的子序列坐标串的公共子序列的点个数。

$$LCSS(i,j)=\begin{cases}0 \quad if \quad i=0 \ or\ j=0\\ LCSS(i-1,j-1)+1\\ \quad if\ \ close\ (a_i,b_i)=true\\ max(LCSS(i,j-1),LCSS(i-1,j))\\ \quad otherwise\end{cases} \tag{1}$$

定义 2　给定一个阈值 ε，函数 close 用来表示两个点是否是同一个点。其定义如下：

$$close(a_i,b_j)=\begin{cases}true\\ \quad |a_{x,i}-b_{x,j}|<\varepsilon,|a_{y,i}-b_{y,i}|<\varepsilon\\ false\\ \quad otherwise\end{cases} \tag{2}$$

定义 3　函数 S 表示两条道路的相似度。

$$S(A,\ B)=\frac{LCSS(m,n)}{\min(m,n)} \tag{3}$$

根据上面的定义，计算两条道路的相似度。本文选择的 S(A, B)的阈值是 0.6。若相似度超过 0.6，就断定两条路基本匹配。若有多个匹配路段，则选取匹配度最高的路段作为命中路段。实际的处理效果与路网的预处理效果、路网特征以及阈值的大小等都有关。此外，需要注意的是，在进行路网匹配之前，需要建立索引。

3.1.3 算法实现

LCSS 算法伪代码如下：

```
算法　LCSS 算法
输入：两个待匹配的坐标串。
输出：LCSS 的 DP 矩阵。
Function getTypeMatrix(coors1,coors2)
m ← coors1.length
n ← coors2.length
type ← new int[m][n]
for i ← m to 1
    for j ← n to 2
        if isClose(coors1[i],coors2[j])
            type[i,j] ← type[i+1,j+1] + 1
        else
            type[i,j]                                    ←
Math.max(type[i][j+1],type[i+1][j])
return type

Function isClose(point1,point2)
x_abs ← |point1.x-point2.x|
y_abs ← |point1.y-point2.y|
if x_abs<distThre&&y_abs<distThre
   return true
else
   return false

Function getMatchRatio( )
ratio ← this.LCS.length/Math.min(L1.length,L2.length)
return ratio
```

上面的 LCSS 算法依赖的数据结构如下：

```
{
  Coordinate[ ] L1;//坐标串 1
  Coordinate[ ] L2;//坐标串 2
  Coordinate[ ] LCS;//经过匹配后的公共坐标串
  double distThre;//经纬度差值的距离阈值
  double matchRatio;//匹配度
  static final double DEFAULT_DISTTHRE=0.0005;//经纬度差值
阈值大约 0.001 在地图上相差 80~90 m
  int commonLen;//公共坐标串的长度
}
```

上述的 Coordinate 包含经度和纬度两个成员变量。

3.2　路网融合

目前城市道路网往往将道路要素分割成多条路段，道路的完整性被破坏。高德路网也是如此，许多道路都经过分段处理，一条语义完整、空间连续的道路可能会被分割为多条路段。因此本文中的路网融合主要指的就是针对这种分段道路进行拼接，使之成为一条平滑无缝的完整道路。

Stroke 方法是道路综合中重要的路段拼接方法，其简单有效的优点正是本文采用该方法的原因。但是传统的 Stroke 方法因为随机选择道路的起始路段，重复实验时会产生不同的 Stroke 拼接结果。因此本文在进行路网融合时，考虑将道路等级作为道路的重要程度并依此来定义路段拼接时的处理顺序，这会保证最终生成的 Stroke 连接结果固定。路网融合时，使用顾及道路语义名和道路等级的 Stroke 连接方法，生成策略采用自身最大合适策略（self-best-fit,SBF）[19]。 所谓 SBF 就是指根据最满足要求的一条路段作为某路段的拼接路段。本文选取的是根据几何规则来判定，即两者拼接时夹角最大。

该算法的主要步骤如下：

a）对路网进行道路重要等级排序，排序结果作为 Stroke 连接的处理顺序。

b)按照顺序进行路网拼接,拼接时先检查道路名是否为空，道路名为空和不为空的处理方式分别按照 c）和 d）进行处理。

c）对于有名路段，如果某一条路段附近有同名路段则直接拼接；如果该路段附近不存在同名路段则不拼接。

d）对于无名路段，则按照“道路等级—最小夹角”的优先顺序进行拼接。所谓的“道路等级—最小夹角”的优先顺序指的是，如果该无名路段附近有相同等级的无名路段且两者夹角超过一定阈值，那就判定为同一道路，否则不拼接。

若不满足上面任何一条处理规则的，则不予处理。上面的处理过程中，如果某路段已经被拼接过，需要做标记，下次进行 Storke 拼接时不再考虑已经标记过的路段。

此外，在路网融合过程中，主要有两种情况需要考虑，即非交叉口道路 Stroke 连接（图 5(a)）和道路交叉口 Stroke 连接（图 5(b)）。

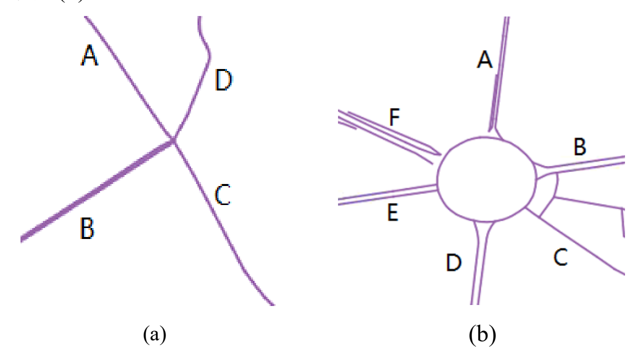


图 5　路网类别示例

3.2.1 非交叉口道路 Stroke 连接

在进行非交叉口的道路的 Stroke 连接时，首先按照道路的重要等级来定义路段的处理顺序，即按照高德路网的 ROAD_CLASS 字段定义路段的处理顺序。对于图 5(a)，该示例只展示了四条路段 A、B、C、D，按照自身最大合适策略（SBF）对图 5(a)进行拼接并不会有什么问题，但是由于 SBF 每处理一条 Stroke 连接时会随机选择起始的处理路段，当路段很多时，最终的 Stroke 处理结果并不固定。因此本文利用道路等级作为处理顺序，改进了 SBF，避免了这个问题。

3.2.2 交叉口 Stroke 连接

对于交叉口的 Stroke 连接，一个重要的前提工作就是交叉口的特征识别。对于图 5(b)这类交叉口，需要判断是否有环状路段，主要方法就是根据路段总共转过的角度来总和是否接近 360°（可能部分路段没有完全衔接上，所以不一定等于 360°）。在进行特征识别之后，再根据 3.2.1 中的方法类似处理，将经过交叉口隔断的路段进行拼接。

4　实验及应用

4.1　路网匹配融合实验

为检验匹配融合技术在道路空间化中的应用效果，本文截取武汉市部分区域的数据作对比，原高德的路网数据如图 6 所示。经过预处理、路网匹配融合之后的路网如图 7 所示。从图中可以看出，多车道的合并、碎段的消除以及其他简化处理使得经过处理后的路网更加简单。



图 6　处理前路网



图 7　处理后路网

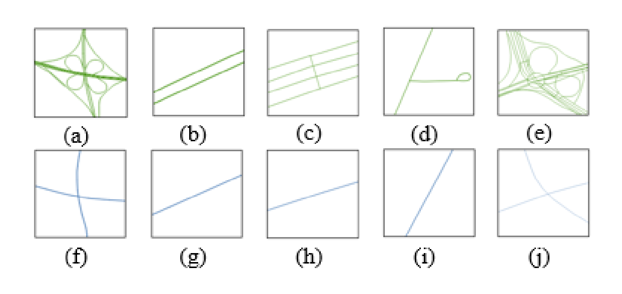


图 8　处理前后对比

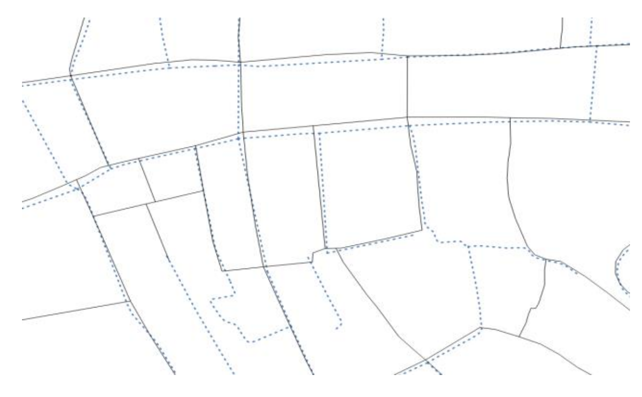


图 9　OSM 路网/简化后的高德路网

针对路网预处理的效果，可以参看图 8。图 8（a）和（f）为高速路匝道消除后的效果；（b）和（g）为多车道合并的效果；

（c）和（h）为消除多车道间碎段的效果；（d）和（i）为消除主干道附近碎段的效果；（e）和（j）为复杂道路的简化效果。

路网匹配的效果如图 9 所示。其中实线为 OSM 路网，虚线为简化后的高德路网。从图中可以看出路网匹配效果较好。

4.2 系统应用

本文基于武汉市道路空间化的需求下，开发了一套路网匹配融合程序。高德路网和 OSM 路网的匹配融合技术在武汉市交通管理局得到了成功地应用。

该程序基于采用了.net com 组件、GDAL 空间库、ArcObject 工具开发包等技术，提供了坐标系转换、道路数据预处理、属性计算、道路简化、路网融合五大组件服务。

原型系统界面如图 10 所示。该系统左侧面板即为图层展示部分，提供了数据添加、缩放、平移、保存等功能；中间部分即为数据展示区；右侧上方为参数设置区；右侧下方即为整个路网处理流程。用户可以基于此系统根据实际处理效果，自定义参数大小。此系统提供了弯曲范围、道路夹角、缓冲距离等三个参数。另外，一键处理路网使得该系统使用简单，极大地方便了用户。

此外，该系统提供的路网匹配融合功能，也在其他系统中得到应用，如图 11 所示。

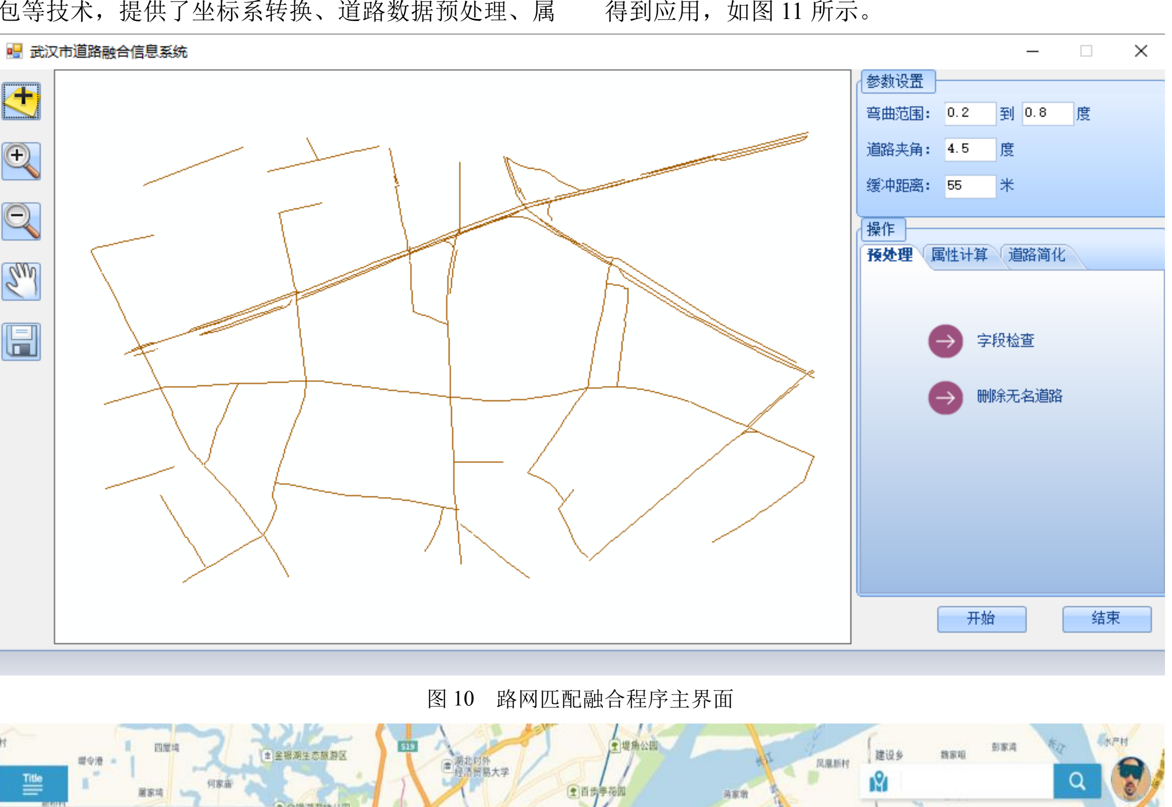


图 10 路网匹配融合程序主界面

图 11 道路空间化系统主界面

5 结束语

目前武汉市交通管理局对于事故、违法定位都是文本描述的方式，随着移动终端 PDA 的普及以及道路管理的更新和维护的需要，空间路网数据的必要性凸显得越来越重要。本文基于这种背景，设计并开发了一套多源路网匹配融合系统，解决了多源路网管理复杂、维护困难、人工操作耗时耗力的问题。同时，该系统已经在武汉市交通管理局上线投入使用，为事故定

位和违法定位提供了路网数据的支撑，给武汉交管部门的信息化和道路空间化作出了重要的贡献。

参考文献：

[1] 刘一宁. 多源矢量道路匹配方法研究 [J]. 现代测绘, 2014 (2): 16-18. (Liu Yining. Research on multi-source vector road data matching [J]. Modern Surveying and Mapping, 2014 (2): 16-18.)

[2] 郭庆胜, 谢育武, 刘纪平, 等. 顾及尺度变化和数据更新的道路网匹配算法 [J]. 测绘学报, 2017, 46 (3): 381-388. (Guo Qingsheng, Xie Yuwu, Liu Jiping, *et al.* Algorithms for road networks matching considering scale variation and data update [J]. Acta Geodaetica et Cartographica Sinica, 2017, 46 (3): 381-388.)

[3] 张云菲, 杨必胜, 栾学晨. 利用概率松弛法的城市路网自动匹配 [J]. 测绘学报, 2012, 41 (6): 933-939. (Zhang Yunfei, Yang Bisheng, Luan Xuechen. Automated matching urban road networks using probabilistic relaxation [J]. Acta Geodaetica et Cartographica Sinica, 2012, 41 (6): 933-939.)

[4] 巩现勇, 武芳, 姬存伟, 等. 道路网匹配的蚁群算法求解模型 [J]. 武汉大学学报: 信息科学版, 2014, 39 (2): 191-195. (Gong Xianyong, Wu Fang, Ji Cunwei, *et al.* Ant colony optimization approach to road network matching [J]. Geomatics and Information Science of Wuhan University: Information Science Edition, 2014, 39 (2): 191-195)

[5] 安晓亚, 孙群, 肖强, 等. 一种形状多级描述方法及在多尺度空间数据几何相似性度量中的应用 [J]. 测绘学报, 2011, 40 (4): 495-501. (An Xiaoya, Sun Qun, Xiao Qiang, *et al.* A shape multilevel description method and application in measuring geometry similarity of multi-scale spatial data [J]. Acta Geodaetica et Cartographica Sinica, 2011, 40 (4): 495-501.)

[6] 安晓亚, 孙群, 尉伯虎. 利用相似性度量的不同比例尺地图数据网状要素匹配算法 [J]. 武汉大学学报: 信息科学版, 2012, 37 (2): 224-228. (An Xiaoya, Sun Qun, Yu Bohu. Feature matching from network data at different scales based on similarity measure [J]. Geomatics and Information Science of Wuhan University: Information Science Edition, 2012, 37 (2): 224-228.)

[7] Vlachos M, Gunopoulos D, Kollios G. Discovering Similar Multidimensional Trajectories [C]// Proc of the 18th International Conference on Data Engineering. Piscataway, NJ: IEEE Press, 2002: 673-684.

[8] 罗路长, 刘波, 刘雪朝. OpenStreetMap 路网数据质量评价及应用分析 [J]. 江西科学, 2017, 35 (1): 151-157. (Luo Luchang, Liu Bo, Liu Xuechao. Data quality evaluation and application analysis of OpenStreetMap road network [J]. Jiangxi Science, 2017, 35 (1): 151-157.)

[9] Mooney P, Corcoran P, Winstanley A C. Towards quality metrics for OpenStreetMap [C]// Proc of the 18th Sigspatial International Conference on Advances in Geographic Information Systems. New York: ACM Press, 2010: 514-517.

[10] Zielstra D, Zipf A. A comparative study of proprietary geodata and volunteered geographic information for Germany [C]// Proc of the 13th AGILE International Conference on Geographic Information Science. Berlin: Springer-Verlag, 2010: 1-15.

[11] Wang Ming, Li Qingquan, Hu Qingwu, *et al.* Quality analysis of open street map data [J]. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2013, XL-2//W1: 155-158.

[12] Zhang Meng, Shi Wei, Meng Liqiu. A generic matching algorithm for line networks of different resolutions [C]// Proc of Ica Workshop on Generalisation & Multiple Representation. 2005: 7-8.

[13] 翟仁健. 基于全局一致性评价的多尺度矢量空间数据匹配方法研究 [D]. 郑州: 解放军信息工程大学, 2011. (Zhai Renjian. Research on automated matching methods for multi-scale vector spatial data based on global consistency evaluation [D]. Zhengzhou: Information Engineering University, 2011.)

[14] Zhang Zhengyou. Iterative point matching for registration of free-form curves and surfaces [J]. International Journal of Computer Vision, 1994, 13 (2): 119-152.

[15] Berndt D J, Clifford J. Using dynamic time warping to find patterns in time series [C]// Proc of the 3rd International Conference on Knowledge Discovery and Data Mining. Seattle, WA: AAAI Press, 1994: 359-370.

[16] Huttenlocher D P, Klanderman G A, Rucklidge W J. Comparing images using the Hausdorff distance [J]. IEEE Trans on Pattern Analysis & Machine Intelligence, 1993, 15 (9): 850-863.

[17] 李欣, 舒风笛. 最长公共子序列问题的改进快速算法 [J]. 计算机应用研究, 2000, 17 (2): 28-30. (Li Xin, Shu Fengdi. Improved fast algorithm for the longest common subsequence problem [J]. Application Research of Computers, 2000, 17 (2): 28-30.)

[18] 刘宇, 王前东. 基于最长公共子序列的非同步相似轨迹判断 [J]. 电讯技术, 2017, 57 (10): 1165-1170. (Liu Ning, Wang Qiandong. Computing similar measure between two asynchronous trajectories based on longest common subsequence method [J]. Telecommunication Engineering, 2017, 57 (10): 1165-1170.)

[19] 田晶, 任畅, 王一恒, 等. 对生成 Stroke 的自身最大适合策略的改进 [J]. 武汉大学学报: 信息科学版, 2015, 40 (9): 1209-1214. (Tian Jing, Ren Chang, Wang Yiheng, *et al.* Improvement of self-best-fit strategy for stroke building [J]. Geomatics and Information Science of Wuhan University: Information Science Edition, 2015, 40 (9): 1209-1214.)